Proceedings of 10th Asian-Pacific Regional IAU Meeting 2008, 1–3 (2008) Printed 9 September 2009 (APRIM LaTeX style file v1.0)

On the Period-Mass Functions of Extrasolar Planets

Li-Chin Yeh[1], Yen-Chang Chang[1], Wen-Liang Hung[2], and Ing-Guey Jiang[3]

[1] *Department of Applied Mathematics, National Hsinchu University of Education, Hsin-Chu, Taiwan*
[2] *Graduate Institute of Computer Science, National Hsinchu University of Education, Hsin-Chu, Taiwan*
[3] *Department of Physics and Institute of Astronomy, National Tsing-Hua University, Hsin-Chu, Taiwan*

ABSTRACT

Using the period and mass data of two hundred and seventy-nine extrasolar planets, we have constructed a coupled period-mass function through the non-parametric approach. This analytic expression of the coupled period-mass function has been obtained for the first time in this field. Moreover, due to a moderate period-mass correlation, the shapes of mass/period functions vary as a function of period/mass. These results of mass and period functions give way to two important implications: (1) *the deficit of massive close-in planets is confirmed*, and (2) *the more massive planets have larger ranges of possible semi-major axes*. These interesting statistical results will provide important clues into the theories of planetary formation.

Key words: planetary systems, extrasolar planets, distribution functions, correlation coefficients

1 INTRODUCTION

After the first detection of an extra-solar planet (exoplanet) around a millisecond pulsar in 1992 (Wolszczan & Frail 1992), it was soon reported that another exoplanet, the first one around a sun-like star, i.e. 51 Pegasi b, was found (Mayor & Queloz 1995). Ever since then, there has been a continuous flood of discoveries of extra-solar planets. As of February 2008, more than 200 planets have been detected around solar type stars. These discoveries have led to a new era in the study of planetary systems. For example, the traditional theory for the formation of the Solar System does not likely explain certain structures of extra-solar planetary systems. This is due to the properties, discovered in extra-solar planetary systems, being quite unlike our own. Many detailed simulations and mechanisms have been proposed to explore these important issues (Jiang & Ip 2001, Kinoshita & Nakai 2001, Armitage et al. 2002, Ji et al. 2003, Jiang & Yeh 2004a, Jiang & Yeh 2004b, Boss 2005, Jiang & Yeh 2007, Rice et al. 2008).

As the number of detected exoplanets keeps increasing, the statistical properties of exoplanets have become more meaningful. For example, assuming that the mass and period distributions are two independent power-law functions, Tabachnik & Tremaine (2002) used the maximum likelihood method to determine the best power-index. However, the possibility of a mass-period correlation is not addressed in their work. Zucker & Mazeh (2002) determined the correlation coefficient between mass and period in logarithmic space and concluded that the mass-period correlation is significant.

On the other hand, a clustering analysis of the data we have on exoplanets also gives some interesting results. Jiang et al. (2006) took a first step into clustering analysis and found that the mass distribution is continuous, and the orbital population could be classified into three clusters which correspond to the exoplanets in the regimes of tidal, ongoing tidal and disc interaction. Marchi (2007) also worked on clustering through different methods.

To take things a step further from the mass-period distribution function of Tabachnik & Tremaine (2002) and the mass-period correlation of Zucker & Mazeh (2002), Jiang, Yeh, Chang, & Hung (2007) (hereafter JYCH07) employed an algorithm to construct a coupled mass-period function numerically. They were able to include the possible correlation of mass and period into the distribution function for the first time in this field and obtained a distribution function that found a correlation to be consistent. In fact, the mass-period distribution obtained by JYCH07 should be called the mass-period *probability density function* (pdf) in statistics. The integral of pdf is then called the *cumulative distribution function* (cdf). We will use the above terms in this paper.

Although JYCH07 successfully constructed the coupled mass-period pdf numerically, due to constraints in the algorithm they employed, they were forced to use the parametric approach of β-distribution on the pdf fitting. The pdf is a basic characteristic describing the behavior of random variables, i.e. mass and period, and is so important that one has to choose the underlying functional form carefully. One possibility to address this problem is to use the nonparametric approach. This is because the nonparametric approach is a

distribution-free inference. That is, an inference that is made without any assumptions regarding the functional form of the underlying distribution. In addition, the most valuable indication of the nonparametric approach is to let the data speak for itself. We therefore see no other reasonable course of action than to use the nonparametric approach in this paper.

Moreover, we still consider the period-mass coupling even while the pdf and cdf are being constructed. In order to make it possible to proceed, we will employ a method called “Copula Modelling” to obtain the coupled pdf and cdf on the period and mass of exoplanets. This method is more general than the one used in JYCH07 so that a nonparametric approach can be used to obtain the coupled pdf. “Copula Modelling” has a long history of development and was too complicated to be used with real data, in practical terms, until Trivedi & Zimmer (2005) clearly demonstrated a standard modelling procedure.

In §2, we briefly describe our data. The estimation of the nonparametric approach is done as in Jiang et al. (2009). The introduction of the method of Copula Modelling, the demonstration of its credibility, and the application on our data of exoplanets are all described in Jiang et al. (2009). The results will be summarized in §3, and the discussions and conclusions are in §4.

2 THE DATA

We took samples of exoplanets from The Extrasolar Planets Encyclopaedia (http:// exoplanet.eu/catalog-all.php), 2008 April 10. Our samples do not include OGLE235-MOA53b, 2M1207b, GQ Lupb, AB Pic b, SCR 1845b, UScoCTIO108b, or SWEEPS-04 because either their mass or their period data was not listed. The outlier, PSR B1620-26b, with a huge period (100 years), is also excluded.

The data of orbital periods is taken directly from the table in The Extrasolar Planets Encyclopaedia. As a result, only the values of projected mass ($m \sin i$) are listed and only a small fraction of exoplanets' inclination angles i are known so we decided to provide two models of planetary mass in this paper. For the “minimum-mass model”, we simply set $\sin i = 1$ for all planetary systems in the data. For the “guess-mass model”, an inclination angle i within the observational constraint is assigned to a planetary system through a random process and the mass is then determined accordingly. In this case, if the inclination angle i is given in The Extrasolar Planets Encyclopaedia for a particular planet, we simply use its value. If there is no mention of observational constraints, the angle i will be randomly chosen between 0° and 90°. Please note that the unit of period is days, and the unit of mass is Jupiter Mass (M_J).

3 RESULTS

Using the Copula Modelling, the estimate of dependence parameter θ is $\hat{\theta} = 2.3826$ for the minimum-mass model (see Jiang et al. 2009 for all related equations). Through the bootstrap algorithm as described in JYCH07 with the number of bootstrap replications $B = 2000$, the standard error of $\hat{\theta}$ is 0.3669. In order to properly understand the dependence parameter θ, we also obtain the 95% bootstrap C.I. for θ, which is $(1.6514, 3.1190)$. For the guess-mass model, the estimate of θ is $\hat{\theta} = 2.4565$ and its 95% bootstrap C.I. is $(1.7282, 3.1633)$.

Furthermore, in order to check the stability of the guess-mass model, we repeat the random process to generate 100 guess-mass models and apply Copula Modelling on them. The average value of $\hat{\theta}$ is 2.9249 with the standard deviation 0.3349. We then employ the interquartile range (Turky 1977) to check for any outliers of $\hat{\theta}$ from these 100 guess-mass models. The interquartile range is the difference between the first quartile Q_1 and the third quartile Q_3, i.e. $IQR = Q_3 - Q_1$. Inner fences are the left and right from the median at a distance of 1.5 times the IQR. Outer fences are at a distance of 3 times the IQR. The values lying between the inner and outer fences are called suspected outliers and those lying beyond the outer fences are called outliers (Hogg & Tanis 2006).

The smallest, first quartile, median, third quartile and largest of these 100 $\hat{\theta}$ values, denoted by Min, Q_1, Me, Q_3, Max, respectively, are $Min = 2.3730$, $Q_1 = 2.6297$, $Me = 2.8833$, $Q_3 = 3.1968$, $Max = 3.5776$. Therefore, $IQR = 0.5671$ and cutoffs for outliers are $Q_3 + 1.5IQR = 4.0475$, $Q_3 + 3IQR = 4.8981$, $Q_1 - 1.5IQR = 1.7791$, $Q_1 - 3IQR = 0.9284$. Furthermore, we find that

$$Q_1 - 1.5IQR < Min < Max < Q_3 + 1.5IQR.$$

Thus, all 100 $\hat{\theta}$ values of the guess-mass model lie within the inner fences. It means that no outliers exist in these 100 values and so the stability of the guess-mass model is confirmed.

For the minimum-mass model, the Spearman rank-order correlation coefficient (Press et al. 1992) is obtained as $\rho_S = 0.3769$. Through Copula Modelling, we also find the estimate of Genets correlation coefficient ρ_G (Jiang et al. 2009, Genets 1987), which is $\hat{\rho}_G = 0.3792$. It is obvious that the Spearman rank-order correlation coefficient $\rho_S = 0.3769$ is very close to $\hat{\rho}_G$. Moreover, the 95% bootstrap C.I. with the number of bootstrap replications $B = 2000$ for ρ_G is $(0.2691, 0.4811)$. For the guess-mass model, we have $\hat{\rho}_G = 0.3899$ with a 95% bootstrap C.I. $(0.2811, 0.4869)$. These results are all consistent and confirm that *there is a positive period-mass correlation for exoplanets.*

4 CONCLUSIONS

Using the data of exoplanets, for the first time in this field we have constructed an analytic coupled period-mass function $f_{(P,M)}(p, m|\theta)$ through a nonparametric approach. Moreover, we calculate the Spearman rank-order correlation coefficient, which gives the same results for linear and logarithmic spaces, and the results in the previous section show that there is a moderate positive period-mass correlation.

In order to comprehend the implication of our results, in Figure 1(a)-(b), we plot $f_{(P,M)}(p, m|\theta)$ with $m = 1, 5, 10, 15 M_J$ (i.e. the period functions given different masses), and also $f_{(P,M)}(p, m|\theta)$ with $p = 1, 50, 100, 150$ days (i.e. the mass functions given different periods) in logarithmic spaces. Note that all curves in Figure 1 are the results of the guess-mass model. For purposes of comparing,

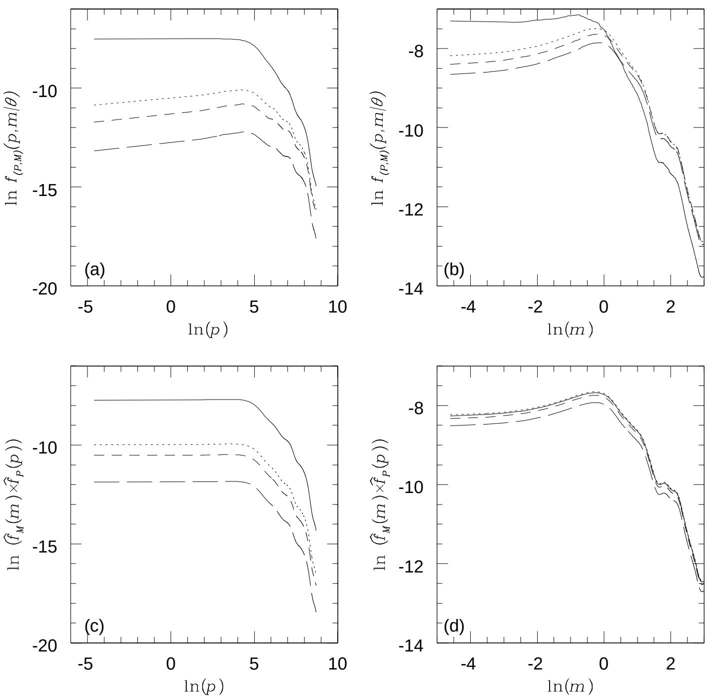

Figure 1. The period and mass functions in logarithmic space. (a) The period functions of $m = 1M_J$ (solid curve), $m = 5M_J$ (dotted curve), $m = 10M_J$ (short dashed curve), and $m = 15M_J$ (long dashed curve). (b) The mass functions of $p = 1$ day (solid curve), $p = 50$ days (dotted curve), $p = 100$ days (short dashed curve), and $p = 150$ days (long dashed curve). (c) The independent period functions of $m = 1M_J$ (solid curve), $m = 5M_J$ (dotted curve), $m = 10M_J$ (short dashed curve), and $m = 15M_J$ (long dashed curve). (d) The independent mass functions of $p = 1$ day (solid curve), $p = 50$ days (dotted curve), $p = 100$ days (short dashed curve), and $p = 150$ days (long dashed curve).

$f_P(p) \times f_M(m)$ with $m = 1, 5, 10, 15M_J$ (the independent period functions) and $f_P(p) \times f_M(m)$ with $p = 1, 50, 100, 150$ days (the independent mass functions) are also plotted in Figure 1(c)-(d). Of course, the shapes of independent period functions with $m = 1, 5, 10, 15M_J$ are all the same, and the shapes of independent mass functions given different periods are all exactly the same as well.

We find that the period function of $m = 1M_J$ is very similar with the independent period functions. However, the period functions of $m = 5, 10, 15M_J$ are different from the independent ones, in a way that the functions are lower at the smaller p end and slightly higher at the larger p end. Thus, the overall period functions of massive planets (say $m = 5, 10, 15M_J$) at large p and small p ends are closer than the one of lighter planets (say $m = 1M_J$). Therefore, the fractions of larger and smaller p (or semi-major-axis) planets are closer for those planets with mass $m = 5, 10, 15M_J$.

This implies that *the more massive planets have larger ranges of possible semi-major axes*. This result is unlikely due to the selection effect because all the planets with masses above $1M_J$ are within the telescopes' detection limits. This interesting statistical result will provide important clues into the theories of planetary formation.

On the other hand, the mass functions of $p = 50, 100, 150$ days are all very similar with the independent mass functions. However, the mass function of $p = 1$ day is different from the independent one in a way that the function is higher at the smaller m end and lower at the larger m end. Thus, the mass function of short period planets (say $p = 1$ day) is steeper than the one of long period planets (say $p = 50, 100, 150$ days). This implies that the percentage of massive planets are relatively small for the short period planets. This result reconfirms *the deficit of massive close-in planets* due to tidal interaction as studied in Jiang et al. (2003).

ACKNOWLEDGMENTS

We are grateful to the National Center for High-performance Computing for computer time and facilities. This work is supported in part by the National Science Council, Taiwan.

REFERENCES

Armitage, P. J., Livio, M., Lubow, S. H., Pringle, J. E., 2002, MNRAS, 334, 248
Boss, A. P. 2005, ApJ, 629, 535
Genets, C. 1987, Biometrika, 74, 549
Hogg, R.V., Tanis, E.A. 2006, *Probability and Statistical Inference* (Pearson Prentice Hall, New Jersey)
Ji, J., Kinoshita, H., Liu, L., Li, G. 2003, ApJ, 585, L139
Jiang, I.-G., Ip, W.-H. 2001, A&A, 367, 943
Jiang, I.-G., Ip, W.-H., Yeh, L.-C., 2003, ApJ, 582, 449
Jiang, I.-G., Yeh, L.-C. 2004a, MNRAS, 355, L29
Jiang, I.-G., Yeh, L.-C. 2004b, International Journal of Bifurcation and Chaos, 14, 3153
Jiang, I.-G., Yeh, L.-C. 2007, ApJ, 656, 534
Jiang, I.-G., Yeh, L.-C., Chang, Y.-C., Hung, W.-L. 2007, AJ, 134, 2061 (JYCH07)
Jiang, I.-G., Yeh, L.-C., Chang, Y.-C., Hung, W.-L. 2009, AJ, accepted
Jiang, I.-G., Yeh, L.-C., Hung, W.-L., Yang, M.-S. 2006, MNRAS, 370, 1379
Kinoshita, H., Nakai, H. 2001, PASJ, 53, L25
Marchi, S. 2007, ApJ, 666, 475
Mayor, M., Queloz, D. 1995, Nature, 378, 355
Press, W. H. et al. 1992, Numerical Recipes in Fortran, Cambridge University Press
Rice, W. K. M., Armitage, P. J., Hogg, D. F. 2008, MNRAS, 384, 1242
Tabachnik, S., Tremaine, S., 2002, MNRAS, 335, 151
Trivedi, P.K., Zimmer, D.M. 2005, Foundations & Trends in Econometrics, 1, 1
Tukey, W.J., 1977, *Exploratory Data Analysis*, Addison-Wesley, Reading, Mass.
Wolszczan, A., Frail, D. A. 1992, Nature, 355, 145
Zucker, S., Mazeh, T., 2002, ApJ, 568, L113